Evaluation of ISLSCP Initiative II satellite-based land cover data sets and assessment of progress in land cover data for global modeling

Eric C. Brown de Colstoun,[1] Ruth S. DeFries,[2] and John R. G. Townshend[2]

[1] As an important component of the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II data collection, eight state-of-the-art land cover/use data sets have been compiled and made consistent with the ISLSCP Initiative II land/water mask in support of global modeling efforts. These data sets contain new and improved global data sets at coarse resolutions (1/4, 1/2 and 1°) describing historical, recent and present land cover conditions and are a testament to the tremendous progress made in this area over the past decade. In addition to the historical data, data describing the subcell heterogeneity in land cover are also provided, both in terms of subcell proportions of land cover classes and vegetation continuous fields such as % tree, grass and bare cover. Here we present the various ISLSCPII land cover data sets and compare the principal satellite-derived data sets and the effect of their respective aggregation methods. We find that despite some notable disagreements among similar classes, the satellite-based data sets agree remarkably well over large portions of the Earth's surface (over 50% for all resolutions). We also find that the methods of aggregation, whether done by a strictly dominant type, or using more information on subcell tree cover, can have an important impact on the final output and need to be considered by the user. Finally, by integrating the vegetation continuous fields data into our analyses we are able to show that the principal differences in terms of discrete land cover classes are in fact transition zones between similar classes.

1. Introduction

[2] Contemporary, accurate, and consistently repeatable global land cover characterizations such as land cover/use classifications and/or vegetation continuous fields (e.g., % tree cover) play an important role in various aspects of global change studies. Present land cover conditions are needed to generate land cover dependent biophysical parameter fields used in many current General Circulation Models (GCMs) and Numerical Weather Prediction Models. These models can simulate atmospheric circulation and climatic variables such as temperature, rainfall, humidity and wind at a fairly coarse spatial scale and under various global warming scenarios [*Dickinson et al.*, 1986; *Sellers et al.*, 1996a]. The inclusion of land into these computerized models has progressed greatly since the first description of land as a "bucket" by *Manabe* [1969]. Most current models are now coupled with Land Surface Parameterization (LSPs) models which have depended on digital compila-tions of global land cover such as those of *Olson et al.* [1983], *Matthews* [1983], *Wilson and Henderson-Sellers* [1985], or more recent satellite-derived land cover maps such as those of *DeFries and Townshend* [1994a], *Loveland and Belward* [1997], *Hansen et al.* [2000], *Friedl et al.* [2002], and *Bartholomé and Belward* [2005], among others. The LSPs come from a realization that vegetation and soils play an important role, both in space and time, in regulating the exchange of energy, gases and water vapor between the biosphere and the atmosphere and, as such, should be included in global simulations [*Charney et al.*, 1975; *Dickinson*, 1983]. The LSPs serve to produce databases or look-up tables of land cover dependent albedo, surface roughness, and evapotranspiration and respiration, parame-ters that control, respectively, the transfer of energy, mo-mentum, mass, and latent and sensible heat between the biosphere and the lower layers of the atmosphere [*Dorman and Sellers*, 1989; *Bonan*, 2002]. These digital land cover maps also provide the means by which to include the fine-scale heterogeneity of land processes within the coarser grid of the GCMs.

[3] Land cover information is also an important input to biogeochemical, ecosystem, and hydrological models which track the cycling of carbon, nutrients, energy and water between the biosphere and the atmosphere [*Melillo et al.*, 1993; *Running and Hunt*, 1993; *Melillo*, 1994; *Nemani and*

[1]Science Systems and Applications, Inc., Biospheric Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

[2]Department of Geography, University of Maryland, College Park, Maryland, USA.

Running, 1996]. Some of these models can simulate the response of terrestrial ecosystems to elevated CO_2 concentrations and/or climate change, for example. By quantifying the gross and Net Primary Production (NPP) of these ecosystems they can help to identify the geographical location of the principal sources and sinks of carbon, and their temporal and spatial variability, as well as providing improved estimates of the size of various global carbon pools [e.g., *Kicklighter et al.*, 1999]. Vegetation type information is important to these models because various plant and tree species have varied mechanisms for photosynthesis and carbon assimilation which can be affected by different stresses, all factors which can in turn significantly alter estimates obtained from the models [*Bonan*, 1995; *Bondeau et al.*, 1999].

[4] Finally, there is still disagreement between estimates of the land cover conversion which has occurred in the past, and is occurring now, but also on the *rates* of change in land cover conversion [e.g., *Skole and Tucker*, 1993] and the impact of such changes on the global carbon cycle [*Plattner et al.*, 2002; *Houghton*, 2003]. Global land cover characterizations can simplify the monitoring of natural or human-induced changes of land cover/use and are important in modeling the consequences of these changes on local and global processes [e.g., *Bonan*, 1997; *Bounoua et al.*, 2002]. Clearly, if land cover can be accurately measured and consistently monitored globally for a period of several years, significant changes over time, and the rates of these changes, could be quantitatively evaluated and some of the above uncertainties potentially reduced. In fact, recent global land cover classifications have been used as a baseline from which land cover change models can be applied to determine the historical land cover change and rates of change, as demonstrated by *Ramankutty and Foley* [1999] for croplands.

[5] While it remains a highly desirable goal to parameterize global models directly from the remotely sensed observations, the use of static or semistatic land cover data sets for model parameterization is still anticipated for the near future. In alignment with the goals of the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II Data Collection, a collection of state-of-the-art land cover data sets has been compiled in collaboration with the data set producers, reprocessed to common spatial resolutions of 1/4, 1/2 and 1° in both latitude and longitude, with common land/ocean boundaries, to support global modeling efforts. The purpose of this paper is to present these ISLSCP Initiative II land cover data sets, compare and contrast several of the satellite-derived data sets, discuss issues of consistency between the various data sets, and provide an assessment of the progress in these data since the publication of the first coarse resolution satellite-based land cover data sets as a part of the first ISLSCP data collection [*Sellers et al.*, 1996b].

2. Background

[6] The single land cover data set provided in the ISLSCP Initiative I data collection was a 1° by 1° land cover map based on satellite data from the Advanced Very High Resolution Radiometer (AVHRR) [*DeFries and Townshend*, 1994a] and was the first such data set to be generated from remotely sensed data at a global scale. The impetus for the production of that data set was a comparison carried out by *DeFries and Townshend* [1994b] of the three most widely available digital global land cover classifications at the time, those of *Olson et al.* [1983], *Matthews* [1983], and *Wilson and Henderson-Sellers* [1985]. *DeFries and Townshend* [1994b] found that only 26% of the total land area was classified as the same land cover type in all three maps and also noted large discrepancies in terms of the spatial distribution of different major land cover types as well as their actual areal extent over the globe. Because satellite remote sensing provides a synoptic view of the Earth and is able to perform consistent and repetitive quantitative measurements of many terrestrial processes at a variety of spatial scales, they argued that remotely sensed data sets could potentially provide the means by which to generate more consistent and accurate global-scale land cover data sets.

2.1. Global Land Cover Maps Compiled From Ground-Based Sources

[7] Historically, land cover classifications have been performed from ground surveys and/or previous maps and the mapping or delineation of land cover types has typically been made by reference to climate, physiognomic characteristics, floristic composition, or geographical location [*Mueller-Dombois*, 1984; *Prentice*, 1990]. Several important points can be made about these classifications. First, they are subjective in that they reflect the biases of the compilers and the variety of sources they depend on. Second, they offer only qualitative information that is not very useful for input to computerized models of global change.

[8] Several digital maps of global vegetation [e.g., *Olson et al.*, 1983; *Matthews*, 1983; *Wilson and Henderson-Sellers*, 1985] have been compiled from a variety of ground-based sources such as paper maps and atlases, and limited satellite data. While the above databases have been used extensively to support climate change studies, they also are influenced by the decisions and choices of the compilers as well as the quality of their sources. As previously noted, these have disagreed both in terms of the land cover present as well as the areal extent of particular biomes [*Matthews*, 1983; *DeFries and Townshend*, 1994b] but, in all fairness, the differences may also reflect the different purposes of each database. Another difficulty in comparing these data is that, because of the varied methods, classification schemes, and age of sources used, it is not always entirely clear whether the maps reflect the potential or actual vegetation cover, except in the case of bioclimatic classifications. Finally, because they have relied on ground-based sources, any updates or changes have been difficult to implement.

2.2. Global Land Cover Characterizations From Remotely Sensed Data

[9] Satellite remote sensing has been, and is currently being, explored as an attractive alternative for *actual* continental to global-scale land cover classifications [*Tucker et al.*, 1985; *Townshend et al.*, 1987; *Loveland et al.*, 1991; *DeFries and Townshend*, 1994a; *Running et al.*, 1995; *Loveland and Belward*, 1997; *DeFries et al.*, 1995, 1998; *Hansen et al.*, 2000; *Friedl et al.*, 2002; *Bartholomé and Belward*, 2005]. These studies used remotely sensed spectral data acquired from instruments such as the AVHRR, the MODerate

Resolution Imaging Spectroradiometer (MODIS) or the Système Probatoire Pour l'Observation de la Terre (SPOT4)-VEGETATION, coupled with their temporal evolution, to separate land cover classes at the continental and/or global scales. These classifications have typically relied on the variability as a function of cover type of the Normalized Difference Vegetation Index (NDVI). This index, defined as the difference of the solar energy reflected from surfaces in the near-infrared and red portions of the electromagnetic spectrum divided by their sum, is recognized as a broad indicator of surface "greenness," photosynthetic activity, and canopy phenology [Asrar et al., 1984; Justice et al., 1985; Daughtry et al., 1992].

[10] The approach of Loveland et al. [1991], Loveland and Belward [1997], and Bartholomé and Belward [2005] is essentially based on utilizing 12 months of NDVI data with an unsupervised classification algorithm. A large database of ancillary information is used as an aid for the human interpretation of the results. Because of the laborious nature of the postprocessing of these unsupervised classification data sets, they have been produced irregularly and are difficult to implement completely objectively and repeatedly for data sets of multiple years. Other techniques, such as those of DeFries et al. [1995, 1998], and Hansen et al. [2000], are supervised classification approaches which rely on a data set of carefully screened global training data derived from Landsat data; they also used the NDVI, but in conjunction with information from the individual spectral bands of AVHRR, including those in the thermal wavelength region, to improve the efficacy of remotely sensed global land cover classifications. The current MODIS land cover algorithm [Friedl et al., 2002] follows the heritage of supervised classification from AVHRR but its inputs are 16-day composites for the individual MODIS land bands and the Enhanced Vegetation Index (EVI) for an entire year. The current land cover algorithm for the Visible/Infrared Imager/Radiometer Suite (VIIRS) scheduled to fly on the future National Polar Orbiting Environmental Satellite System (NPOESS) [Brown de Colstoun et al., 2000] follows from the AVHRR and MODIS heritage as a supervised classification and also uses a decision tree classifier but is closer to the approach used by Hansen et al. [2000] in terms of data inputs. Finally, the same approach used by DeFries et al. [1995, 1998] and Hansen et al. [2000] to generate global land cover products has also been used with linear mixture models and regression tree algorithms to generate global fields of continuous vegetation characteristics such as tree, herbaceous and bare cover [DeFries et al., 1999; Hansen et al., 2002]. These products are found to more closely represent natural gradients and ecotones in vegetation characteristics, as opposed to the classification of cover types into discrete values and in fact may potentially be more useful to global modelers than stratifications by land cover because they scale linearly to coarser resolutions.

[11] The production of these global land cover data sets would simply not have been possible without the production of the input data necessary for the classifications as well as the production of global training data sets from Landsat data [DeFries et al., 1998]. The first data sets of DeFries et al. [1995, 1998] were produced using the Pathfinder AVHRR Land (PAL) data sets at 8 km spatial resolution [James and Kalluri, 1994]. In parallel, and under the auspices of the

Data and Information System of the International Geosphere Biosphere Programme (IGBP-DIS), a global 1 km data set from AVHRR data was produced spanning the years 1992–1996 to address the needs of several of the IGBP's programs [Townshend et al., 1994]. This 1 km data set [Eidenshink and Faundeen, 1994] forms the core input data for several land cover data sets provided in this ISLSCP Initiative II data collection (Table 1), including the University of Maryland (UMD) land cover data set [Hansen et al., 2000], the IGBP-DIScover vegetation classification [Loveland and Belward, 1997], and the UMD continuous fields of vegetation cover [DeFries et al., 2000]. A recent (2000–2001) MODIS land cover product from MODIS collection 4 [Friedl et al., 2002] has also been added to the ISLSCP Initiative II land cover "suite" to provide a linkage to future data sets that will become available with MODIS and VIIRS. We note that a recent global land cover data set based on SPOT-VGT data named GLC-2000 has been produced under the coordination of the European Commission's Joint Research Centre [Bartholomé and Belward, 2005] but was not available within the time constraints for publication in the ISLSCP Initiative II collection.

2.3. ISLSCP Initiative II Land Cover Data Sets

[12] While the ISLSCP Initiative I collection contained a single global land cover data set, the Initiative II collection now contains 8 different state-of-the-art data sets dealing with various aspects of land cover and/or land use (Table 1), including two historical land cover data sets: the historical croplands fractional cover data set of Ramankutty and Foley [1999], covering the period from 1700–1992, and a related historical land cover and land use (1700–1990) data set from the National Institute of Public Health and the Environment (RIVM) in the Netherlands [Klein Goldewijk, 2001]. Ramankutty and Foley [1998] derived a spatially explicit data set of croplands for the year 1992 by synthesizing remotely sensed land cover data (IGBP-DIScover data set in Table 1) with contemporary land inventory data. Furthermore, Ramankutty and Foley [1999] extended this data set back to 1700 using historical land inventory data. By extending their data set back in time, they were also able to produce a land cover map of "potential" vegetation, or the natural vegetation before human alteration or other types of disturbance, which is also included in this collection. Klein Goldewijk [2001] used historical statistical inventories on agricultural land (census data, tax records, land surveys, etc.) and different spatial analysis techniques to create a geographically explicit data set of land use change, with a regular time interval (see Table 1). These two new global data sets of historical land cover change compare fairly well over most of the Earth despite the different modeling approaches and input data used [Klein Goldewijk and Ramankutty, 2004].

[13] Another interesting addition to the ISLSCP II collection is the data set of Still et al. [2003] which identifies the fraction of each cell with a C_4 dominant photosynthetic pathway. This data set was actually produced from various data sets which are also included in this collection: vegetation continuous fields data [DeFries et al., 2000, see Table 1], that describe the percent of a grid cell covered by herbaceous and/or woody vegetation; the historical croplands data set of Ramankutty and Foley [1999]; climate data from

Table 1. Listing of Land Cover Related Data Sets Provided in the International Satellite Land Surface Climatology Project (ISLSCP) Initiative II Data Collection[a]

| Data Category and Data Set Title	Author(s) and Originating Institution	Input Data Temporal Coverage	Spatial Scale	Data Set Comments
Vegetation				
C4 vegetation percentage	Chris Still, University of California at Santa Barbara	1996–1998	1°	% of each cell which possesses the C4 photosynthetic pathway
Continuous fields of vegetation cover	Ruth DeFries, University of Maryland; Matt Hansen, South Dakota State University	1992–1993	1, 0.5, and 0.25°	% tree, grass and bare cover and % needleleaf, broadleaf, deciduous, evergreen for tree cover
Historical croplands fractional cover	Navin Ramankutty and Jonathan Foley, University of Wisconsin	1700–1992	1 and 0.5°	every 50 years (1700–1850); every 10 years (1850–1980); every year (1986–1992)
Historical land cover and land use	Kees Klein Goldewijk, National Institute of Public Health and the Environment (RIVM), The Netherlands	1700–1990	1 and 0.5°	every 50 years (1700–1950); every 10 years (1950–1990);
MODIS land cover product	Mark Friedl, Alan Strahler, John Hodges, Boston University	2000	1, 0.5, and 0.25°	dominant land cover type, fraction of each cover type and classifier confidence for each cell
Potential vegetation	Navin Ramankutty and Jonathan Foley, University of Wisconsin	N/A	1 and 0.5°	represents natural vegetation before human alteration
UMD land cover classification	Matt Hansen, South Dakota State University; Ruth DeFries, University of Maryland	1992–1993	1, 0.5, and 0.25°	dominant land cover type and fraction of each cover type in each cell
Vegetation classification (IGBP-DIScover)	Tom Loveland and Stephen Howard, National Center for EROS (USGS)	1992–1993	1, 0.5, and 0.25°	dominant type and fraction of each cover type; three classification schemes (IGBP, SiB, BATS)
Ancillary				
Land/water masks, land outline overlays, latitude and longitude grids	Tom Logan, Jet Propulsion Laboratory; ISLSCP II Staff	N/A	1, 0.5, and 0.25°	binary water masks and fractional water/land cover in each cell

[a]The data collection can be accessed online at http://www.daac.ornl.gov.

the Climate Research Unit (CRU) at the University of East Anglia in the United Kingdom [*New et al.*, 2000]; and national crop type harvest area statistics from the Food and Agricultural Organization (FAO) of the United Nations (UN-FAO) and the United Stated Department of Agriculture (USDA). We refer the reader and future users to the data set documentation and/or the above references for more in-depth and specific information on the production of the data sets in Table 1.

[14] The AVHHR-based data sets on Table 1 (UMD land cover and continuous fields, IGBP-DIScover) were produced at a native 1 km spatial resolution from a 1 km global AVHRR data set for 1992–1993 [*Eidenshink and Faundeen*, 1994]. In the process of aggregating from 1 km to 1/4, 1/2 and 1° spatial resolutions the percentages of each class within the coarser cell are calculated and allow the dominant land cover type to be determined. Thus each data set contains one global layer with the dominant type and one layer each per cover type showing the percentage of that cover type in each cell (Figure 1). So in addition to improved classification algorithms, improved input data and spatial resolution, the ISLSCP Initiative II land cover "suite" now provides the user with a thorough description of the subcell variability in land cover that was not available in ISLSCP I data. Therefore the users may now also use

their own rules for aggregation using the layers for each land cover class and potentially produce products that better suit their needs.

[15] The IGBP-DIScover product is the only currently available global product that has been validated against a truly global, and statistically valid, independent set of high-resolution data (~70% accuracy), although it is difficult to say what those numbers may correspond to at the 1/4, 1/2, or 1° spatial scales since the validation was done at 1 km resolution [*Scepan*, 1999]. In fact, an analysis by *DeFries and Los* [1999] suggested that the implications of the classification errors for model parameterizations were substantially lessened at coarser resolutions when compared to the native 1 km resolution of the IGBP-DIScover product. Both the UMD and MODIS data sets have been evaluated against subsets of the same data use to train the classifiers and also have global accuracies near 70% [*Hansen et al.*, 2000; http://geography.bu.edu/landcover/userguidelc/consistent.htm]. In addition, the MODIS land cover product provides gridded estimates of classifier confidences for each cell, addressing an additional user-stated need for accuracy estimates of the products. Finally, in this collection, the IGBP-DIScover data set is provided in three different classification schemes (IGBP, SiB, BATS) to better support the needs of the modeling community.
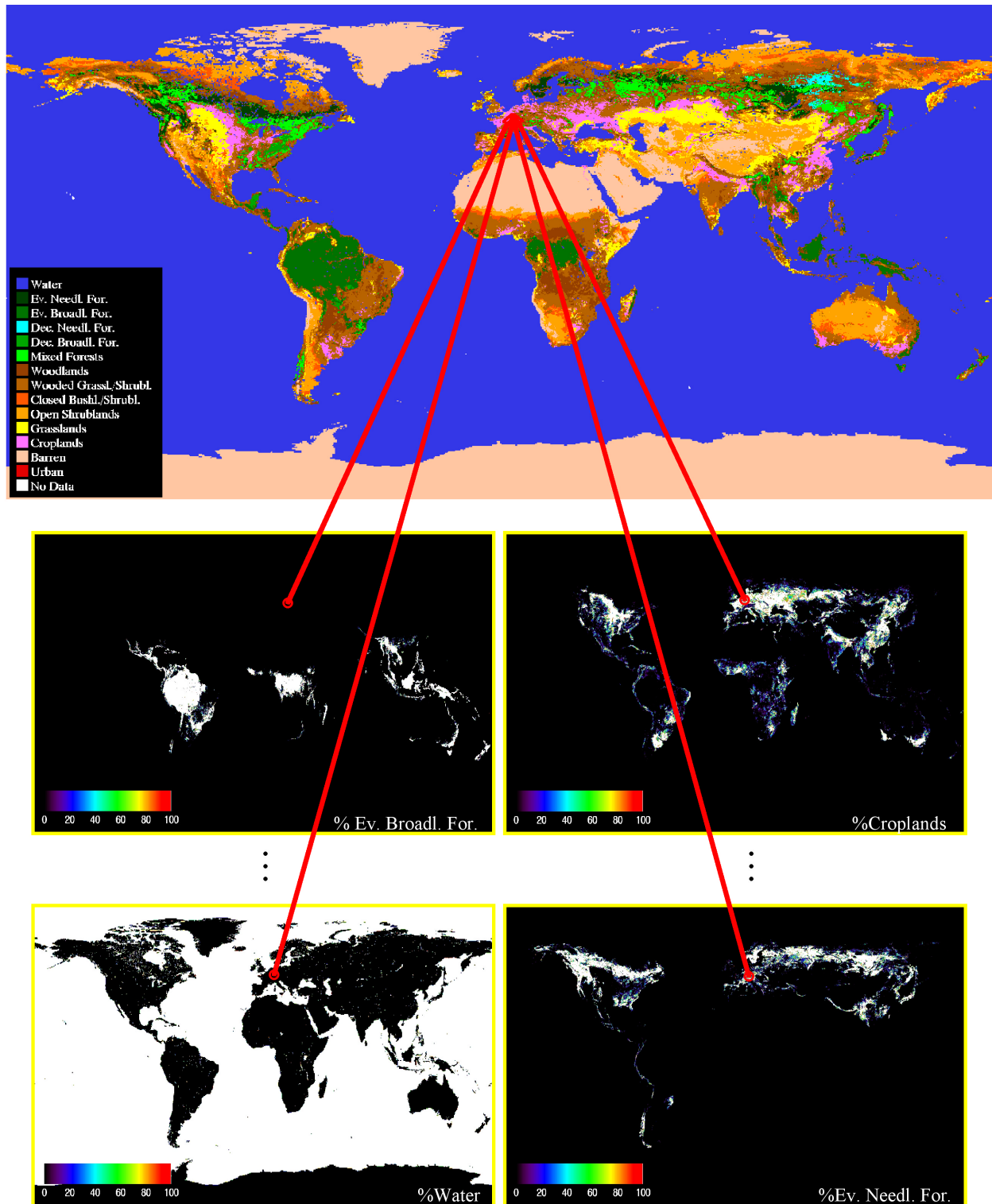
Figure 1. Dominant land cover type determined from the proportions of the individual land cover types at the original resolution of the products. The example here is for the University of Maryland (UMD) land cover product at 1/4° resolution. Only 4 of the 15 layers with subcell percentages of land cover types are shown for clarity. The dominant land cover type in each cell in the top map is determined from a combination of the percentage of each cover type within that same cell.
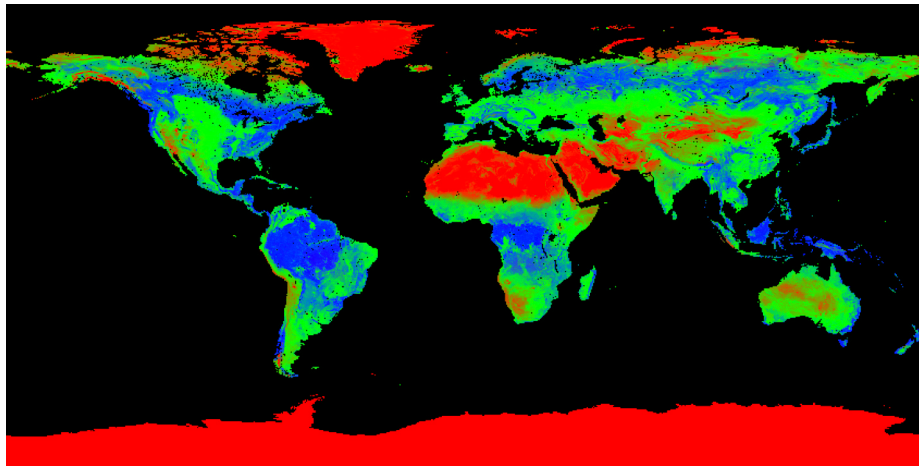
Figure 2. UMD vegetation continuous fields product at 1/4° spatial resolution. The % bare, herbaceous, and woody cover for each cell has been coded as red, green, and blue, respectively, to create this global representation. Other vegetation continuous fields provided in ISLSCP Initiative II include leaf type (needleleaf, broadleaf) and longevity (deciduous, broadleaf) for tree cover.

[16] New types of previously unavailable land cover information are also available at the coarse scale in the form of continuous fields of vegetation cover which describe the % tree, grass and bare cover of each cell (Figure 2), and the % leaf type and/or leaf longevity for tree canopies. All these data sets have been made consistent with the ISLSCP Initiative II land/water mask, which is also based on 1 km original data, and which also contains subcell fractions of water and land at each resolution (see Table 1). This was done by first adjusting the percentage of land and water of each product to correspond to the percentages in the ISLSCP II mask, then recalculating the proportion of each land cover type in the cell on the basis of the new percentage of land, and then producing the dominant cover type maps using aggregation rules specific to each product.

2.4. Previous Land Cover Comparisons

[17] In addition to the product comparisons described by *Matthews* [1983] and *DeFries and Townshend* [1994b], *Hansen and Reed* [2000] have compared the UMD and IGBP-DIScover classifications which are derived from the common 1992–1993 AVHRR data set. They found that the overall per-cell agreement of the two data sets at their original 1 km resolution for all common classes was 48%. For aggregated classes such as forest/woodland, grass/shrubs, crops, this increased to 74% and to 84% when considering even broader classes such as tall woody land cover versus short and/or sparsely vegetated lands. While they found that in general the IGBP-DIScover had more areas of all forest types and the UMD data set showed more areas with intermediate tree cover such as woody savannas and savannas (i.e., woodlands/wooded grasslands), they also found that the principal differences were along transition zones between large core areas. Another difficulty in the comparison was the lack of natural vegetation/croplands mosaic classes in the UMD map, and to a lesser extent, permanent wetlands and ice classes. Finally, *Hansen and Reed* [2000] determined that the overall agreement between IGBP-DIScover and UMD was much greater at 0.5° reso-

lution than the agreement of the well-known digital land cover maps of *Olson et al.* [1983] and *Matthews* [1983]. Their results show a significant decrease of 46% in the amount of disagreement between the remotely sensed data sets as opposed to the digital maps for four broad land cover categories. In this study we have performed similar comparisons using the coarse-scale data in ISLSCP Initiative II at 1/4, 1/2 and 1° resolutions, but we have also analyzed the effects of the aggregation methods on the agreement of the two data sets and used the new data layers available in ISLSCP II (subcell proportion of classes and continuous fields products) to assess the areas of disagreement.

3. Data and Methods

[18] The data sets used in the land cover comparison here were the IGBP-DIScover data set using the 17-class IGBP legend [*Loveland and Belward*, 1997], and the UMD land cover classification using 15 classes [*Hansen et al.*, 2000], at 1/4, 1/2 and 1° spatial resolutions. The MODIS product was not considered in this comparison because of the nearly 10-year gap between the products. The 1/4° UMD vegetation continuous fields product was also evaluated against the 1/4° UMD land cover product to check for data set consistency. As discussed by *Hall et al.* [2006], the incompatibilities in land cover legends between the UMD and IGBP schemes meant that we could only compare similar classes and not classes such as the IGBP natural vegetation/croplands mosaic, which is not included in the UMD product, or the permanent wetlands category. We did however examine the IGBP mosaic class in terms of its subcell makeup in terms of the UMD classes. Also, even though the UMD product did not contain a permanent ice category, we did compare the IGBP-DIScover ice category against the UMD bare category. In total, we were able to compare the two products over 94% of the land surface.

[19] The two data sets have been aggregated to coarser resolution using somewhat different rules that can influence the dominant type found on the final land cover map. The IGBP-DIScover product was aggregated using a strictly
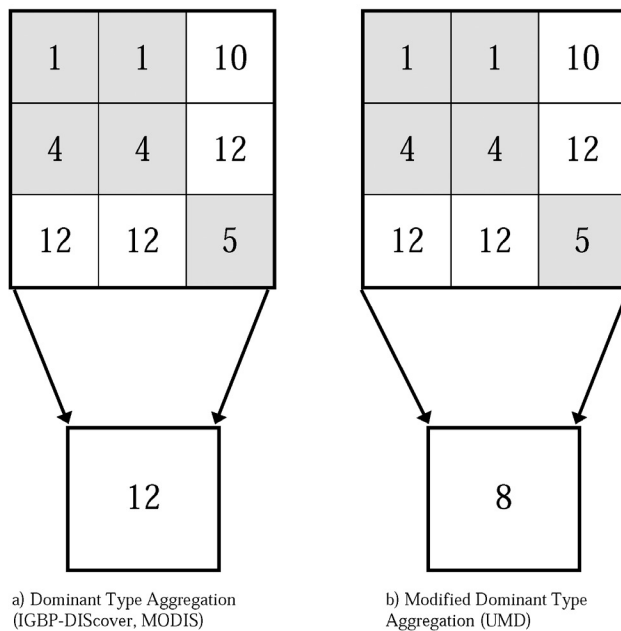
Figure 3. Illustration of two methods of aggregation of finer-resolution land cover products to coarser resolutions. (a) For the IGBP-DIScover and MODIS products, the dominant type is selected from the maximum percentage of any cover type in the cell, irrespective of type. (b) In the UMD "modified" approach, the presence of multiple forest types within the cell is accounted for as well as an estimation of the amount of woody cover. The resulting aggregation is more robust for Figure 3b because the cell is at least 56% forest and only 33% cropland. IGBP classes shown are evergreen needleleaf forest (1), deciduous broadleaf forest (4), mixed forest (5), wooded savannas (8), grasslands (10), and croplands (12).

dominant land cover type, whereby the land cover type with the largest percentage in the cell was selected as the dominant type, regardless of any other information included the cell (see Figure 3a). As can be seen in Figure 3a, however, a purely dominant type approach can create results that may emphasize the importance of single land cover types (e.g., croplands in Figure 3a) over multiple land cover types such as forests. In Figure 3a, croplands is selected as the dominant type even though the three forest types account for 56% of the cell coverage. We should note that Figure 3 shows just an illustration of the aggregation with only nine cells and that for the ISLSCP Initiative II data sets, windows of 120×120, 60×60 and 30×30 cells of 1 km (i.e., 30 arc-seconds) were averaged to obtain the percentages in each 1, 1/2 and 1/4° cell, respectively.

[20] The UMD product was aggregated by the data providers using what we have termed a "modified" dominant type approach (Figure 3b). This approach uses the UMD class definitions, particularly in terms of woody cover, and voting rules that account for the woody cover of the aggregated cell, to assign a dominant type and attempts to overcome some of the issues seen with a strictly dominant approach. For example, this approach assigns forest land cover types when the forest cover of the cell is

greater than or equal to 60%, wooded grasslands for forest covers between 40% and 60%, and so on. From the results shown in Figure 3b, it appears that the modified approach may account for the subcell variability of cover types in a more robust fashion than the dominant approach, although the results globally are not substantially different for most core areas.

[21] We have compared the IGBP-DIScover product with the UMD data set first using a dominant approach for both, and then the modified approach for UMD as is currently provided in the collection, and analyzed the results in terms of overall agreement at the various resolutions, global land cover proportions for each land cover type as well as per class agreements. For the principal areas of disagreement between the two data sets, we have used the subpixel proportions to better understand the nature of the differences, but also checked the makeup of the IGBP natural vegetation/croplands mosaic against the UMD proportions. Finally, we have also used the UMD vegetation continuous fields data sets to both check the areas of disagreement between the two land cover maps but also to check the correspondence and consistency of this product and the UMD land cover product. Clearly, these types of analyses would not have been possible if the data sets did not consistently overlay each other in terms of land/water areas so that the work of making these data sets consistent with the ISLSCP II land/water mask was essential.

4. Results and Discussion

[22] Figure 4 shows the results of the per-pixel comparisons of the two land cover data sets at multiple resolutions. It is important to note that these comparisons provide a good indication as to the level of consistency between the data products and are not meant to imply that one product is necessarily superior. The results show that, when using a strictly dominant criterion for both data sets, the agreement increases as the resolution gets coarser, from 48% at 1 km [*Hansen and Reed*, 2000], to 50.23% at 1/4°, 51.10% at 1/2° and 51.63% at 1°. When adding the agreement of the IGBP permanent ice and the UMD bare categories, the two data sets agree over 68% of the global land surface, a remarkable agreement given the significant algorithmic differences in generating the data sets. In contrast, when comparing the IGBP-DIScover with the modified dominant type UMD map, we find that the agreement decreases with coarser resolution, from 46.87% at 1/4° to 45.55% at 1°. Clearly, it appears that the methods of aggregation to coarser resolution do have an important effect and that the agreement of the data sets increases with coarser resolution for both data sets aggregated using the same dominant method. The modified approach tends to create more and more "mixed" pixels classes such as woodlands as the resolution gets coarser and thus begins to diverge with the strictly modified version which will emphasize the dominance of single cover types. Nonetheless, when looking at the spatial differences between the two maps, our findings are in agreement with those of *Hansen and Reed* [2000] in that the large core areas of land cover are mapped similarly in both products, and the larger differences are found in transition zones between similar cover types. Again the overall agreement of nearly 50% is a clear improvement
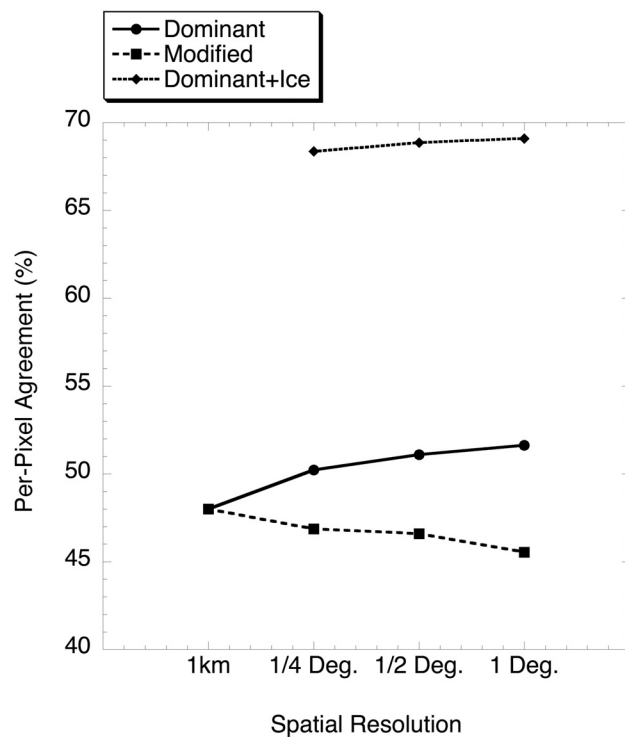
Figure 4. Per-cell agreement of IGBP-DIScover and UMD land cover products at several spatial resolutions and using either a dominant or "modified" dominant aggregation scheme. Results of a comparison including the IGBP-DIScover permanent ice category against the UMD bare category are also shown. The agreement value of 48% at 1 km resolution is from *Hansen and Reed* [2000].

over the 26% agreement found by *DeFries and Townshend* [1994a] using the *Olson et al.* [1983], *Matthews* [1983] and *Wilson and Henderson-Sellers* [1985] maps.

[23] We averaged the subcell proportion layers associated with each map to obtain the percentage of each land cover type on the surface of the Earth estimated by each map, as shown on Figure 5 for the 1° resolution data. Similar to what *Hansen and Reed* [2000] did with global land area totals for each class but with proportions here, we find that at this resolution the IGBP-DIScover map contain more forests than the UMD map (+6.51%) while the UMD has substantially more woody savannas and savannas (+13%), corresponding to the woodlands and wooded grasslands UMD classes. Proportions for the shrublands classes, grasslands and croplands are quite close, within 2%, while the proportion of the IGBP ice and bare categories are 1% higher than the UMD bare categories. We also found that the proportions for the urban classes of the two data sets did not match, even though both data sets used the same Digital Chart of the World as the source for this class. Differences in the water proportions for both data sets were also found but addressed by making the data consistent with the ISLSCP II land/water mask.

[24] An examination of the typical areas of disagreement for each land cover class showed substantial disagreement between the five different forest types, although the evergreen broadleaf forest compared quite well between the two data sets. Over 90% of this class in the UMD map was mapped accordingly in the IGBP-DIS while 72.28% of all the IGBP-DIScover evergreen broadleaf forest was mapped as such in the UMD data set. In general the UMD forest types, except for deciduous broadleaf forest, were fairly well mapped in the IGBP-DIScover product. Reasonable agreement was found for the open shrublands (>50%), croplands, grasslands and bare categories. Fairly large disagreement was found in the woody savanna, savanna, and closed shrublands classes but very often the source of this disagreement was between similar classes such as
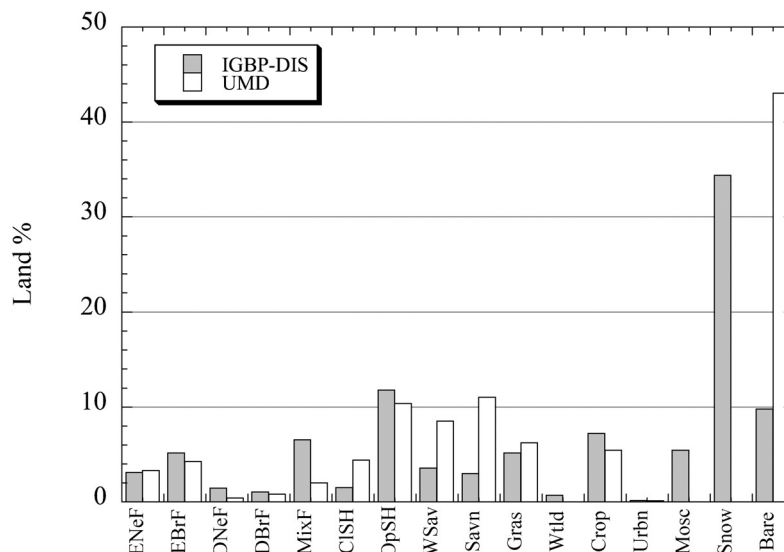
Figure 5. Global land cover proportions for the 1° IGBP-DIScover and UMD land cover products derived from an average of the subcell proportion layers at the same resolution.
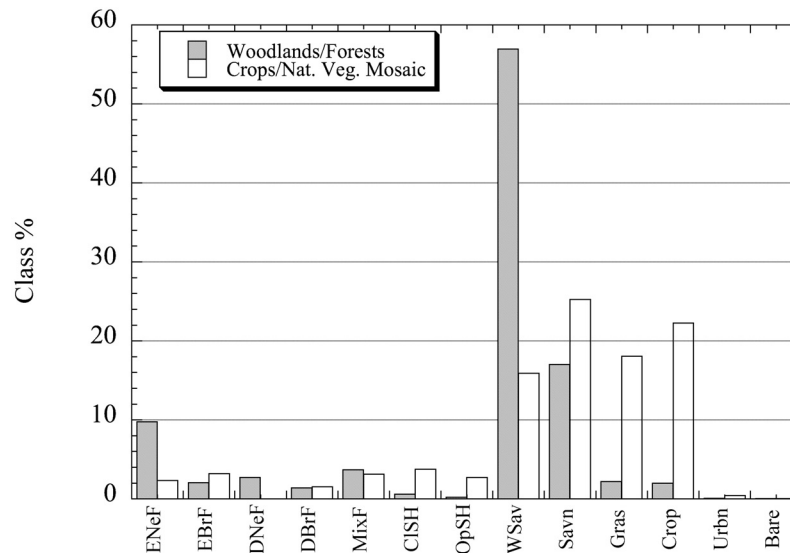
Figure 6. Subcell composition of areas of disagreement between the IGBP-DIScover five forest classes and the UMD woodlands and woody grasslands classes. The graph also shows the subcell composition of the IGBP-DIScover natural vegetation/croplands mosaic class in terms of UMD cover types. These are the types of analyses and comparisons that can be made with the subcell proportion data available in this collection.

woody savanna and savanna, or between closed and open shrublands, for example, and not between core classes like forests and bare soil or croplands. In fact, 15% of the total disagreement could be attributed to the UMD woodlands (i.e., woody savanna) class alone, and together with the wooded grasslands (i.e., savanna) class, accounted for 22% of all the total disagreement. Likewise, the IGBP-DIScover mixed forest class was confused across almost all the UMD classes. When grouping all forest classes into one, both savanna classes into another and both shrublands classes into yet another, the agreement was approximately 60%, indicating that core classes compared well among the two products, as shown by *Hansen and Reed* [2000].

[25] We used the tools provided in the ISLSCP Initiative II collection to assess the typical areas of disagreement between the data sets and also as a way to explore the "severity" of these disagreements. As an example, we used the per-pixel proportions to check the composition of the most common areas of disagreement which were between the UMD woodlands and wooded grasslands classes and the IGBP-DIScover forest classes. Results are shown in Figure 6. What Figure 6 shows is that for the most part the areas of disagreement are made up by a majority of the woody savanna class, about 17% of the savanna and 10% evergreen needleleaf forest. What this analysis confirms is that the disagreements are indeed between classes with similar tree cover and are found in areas where forests transition into more open canopies and then into savannas, such as the ecotones of the boreal forest, or the Miombo woodlands of Africa. In Figure 6 we also show the subcell composition of the IGBP-DIScover natural vegetation/croplands mosaic class in terms of the UMD cover types which show that this class is made up of a mixture of croplands with the other cover types, but principally with the savanna (i.e., woody grasslands), grasslands and woody

savannas (i.e., woodlands). This is entirely consistent with the definition of this IGBP mosaic class and show how these new types of data available in this collection can be used for meaningful analyses.

[26] The areas of disagreement have also been explored with the UMD vegetation continuous fields data. Although this data set was generated separately from the UMD land cover product and the approaches to generate each product are different, we should note that the input AVHRR data and the training data are the same for both data sets. Figures 7a–7c show histograms for % grass, % tree and % bare cover for each cell in the 1/4° data sets where we found disagreement between the IGBP-DIScover forest types and the UMD woodlands classes. For all of these cells the mean value of % bare, grass and tree cover was 2.76, 51.21, and 44.75%, respectively. These figures show that, while there are some apparent inconsistencies between the two UMD products with unusually high and/or low values of tree cover or grass cover, and also while there are some substantial disagreements between the two land cover maps, this disagreement is not as significant when considered in terms of tree cover. Figure 7b shows that indeed many of the confused areas are likely to be transition zones between forest types, with greater than 60% tree cover, to woody savannas with 40 to 60% tree cover, as seen in Figure 6 as well. Figure 7a also confirms the confusion with savanna classes with tree cover values between 10 and 40% and Figure 7c confirms that there are few areas of confusion with areas of low tree cover. These percentages of tree cover are consistent with the IGBP-DIScover and UMD definitions of woody savanna and savanna classes yet the smaller differences seen here in terms of tree cover are amplified when comparing discrete classes. We would like to reemphasize to potential users that while the maps may not always agree on a category by category basis and cell by

cell basis, the differences in terms of the actual canopy cover types considered are usually not large.

[27] As a final step in our evaluation, we have compared the UMD continuous fields and land cover products to check for internal consistency and also as an evaluation of the UMD continuous fields data provided here. Table 2 shows the mean continuous fields values for all cells in each UMD land cover type from the 1/4° data sets. Overall, the mean proportions of bare, grass and tree cover are entirely consistent with the UMD land cover definitions for each land cover type, with forest types above 60% tree cover, for example, woodlands with >40% tree cover and woody grasslands with tree cover greater than 10%. There are also some outliers where we find bare pixels with high tree cover

Table 2. Mean Subpixel Proportions for Each UMD Land Cover Type in Terms of the UMD Vegetation Continuous Fields Data

Land Cover Class	% Bare	% Grass	% Tree
ENeF	1.79	35.4	60.98
EBrF	2.60	22.11	73.64
DNeF	1.78	35.63	61.74
DBrF	1.11	30.01	67.91
MixF	1.73	31.02	65.91
ClSh	15.58	79.68	3.42
OpSh	38.15	59.65	0.69
Wdld	2.91	53.86	41.95
WoGr	4.12	72.32	22.09
Gras	4.12	87.50	4.61
Crop	1.87	87.61	9.5
Urbn	2.41	76.36	15.77
Bare	95.47	3.83	0.02
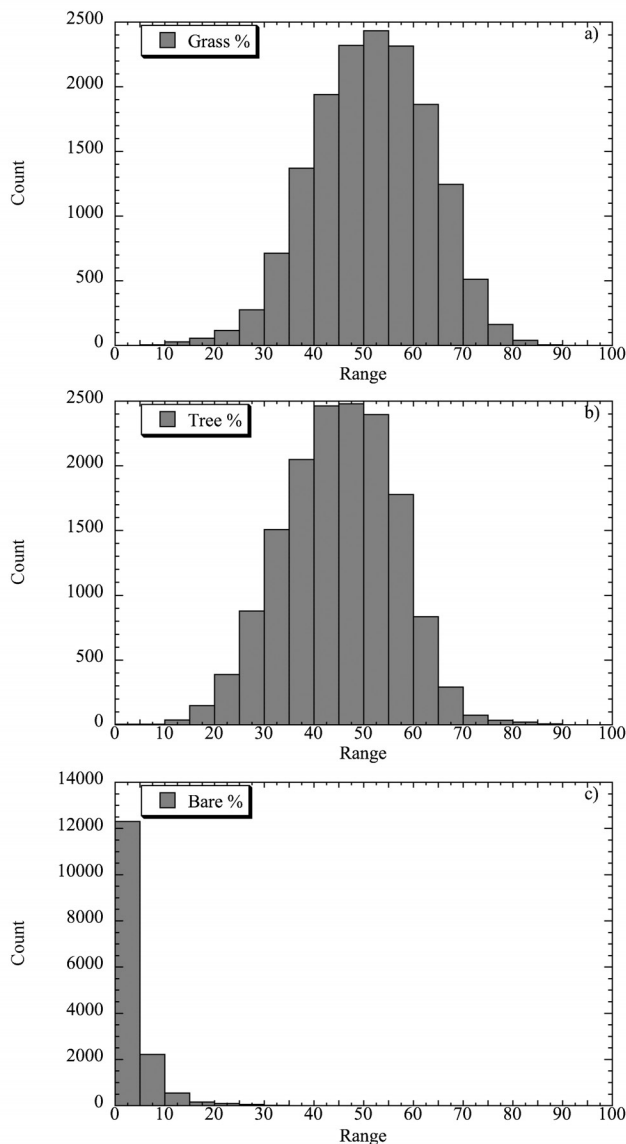
and/or forest types with high bare cover, but these are uncommon. These inconsistencies do point to the need for a potentially common, or at least internally consistent, processing method(s) for these types of data sets. Finally, the mean values in Table 2 for both the evergreen needleleaf and deciduous needleleaf forest classes are lower than for the other classes, and close to 60% tree cover, which may explain some of the disagreements seen with the IGBP-DIScover maps for these classes. Likewise the mean value of 41.95% tree cover for the woody savannas class is somewhat low and may be indicative of the source of disagreements with this class. Also, it is interesting to note that the mean composition of the urban class according to these data contains very little bare areas but again this class has not been provided by the remotely sensed data but rather has been superimposed from a static database.

5. Conclusions

[28] The various land cover data sets provided in the ISLSCP Initiative II data collection to support global modeling efforts represent the tremendous progress made in this area over the past decade or so. From the first global data set from remotely sensed data provided in ISLSCP I [*DeFries and Townshend*, 1994a], we have now progressed to multiple and improved data sets that accurately describe the past, present and future land cover conditions on the Earth. The algorithms to generate such data sets have progressed to include machine learning classifiers such as decision trees [*Hansen et al.*, 2000; *Friedl et al.*, 2002] which can efficiently handle nonnormal distributions in the training data; new global training data sets generated from high-resolution data and improved input data at higher spatial resolutions are now available and being used for validation as well; the number and types of classes that can be provided has been greatly expanded, and finally, particularly for modeling applications, the subcell variability in terms of land cover types is provided with each data set. New types of data sets with subcell information have also become available such as the vegetation continuous fields data and the C_4 fraction data set of *Still et al.* [2003], giving the user great flexibility in land cover class definitions but also providing a better representation of landscape continuity across land cover types and ecotones. All of these new and improved data sets should in turn provide improved

Figure 7. Histograms showing the subcell proportions of (a) % grass, (b) % tree, and (c) % bare cover from the UMD vegetation continuous fields data set for all cells where the IGBP-DIScover forest types and the UMD woodlands categories disagreed at a 1/4° resolution.

estimates for those models that are using them for land surface parameterizations.

[29] Our comparison of the two most widely used land cover data sets for the 1990s shows that, despite some real differences at the level of individual classes, the two data sets agree over nearly two thirds of the Earth's surface and for large core classes. The comparison also shows that clearly, the methods used for aggregation of the products can have a significant impact on the final land cover product. However, while the "modified dominant" approach does appear to produce more robust results, there are to date no guidelines or "best practices" for users to follow in the aggregation of land cover from moderate to coarse resolutions. Also, more analyses are needed to assess the impact of these aggregation methods on the final results, and their subsequent impacts upon the models that use them.

[30] The areas of disagreement, when considered in terms of tree cover, are shown to be transition zones between similar classes and as such their impact on modeling studies may not be as severe as the disagreement between discrete classes seems to indicate. The level of agreement between the data sets also shows a marked improvement over the agreement of previously available digital data sets. However, the data set inconsistencies seen here do point to the need for better integration and harmonization of efforts and more consistent approaches aimed at reducing interproduct differences and thus facilitate the use of the data. It is also critical to note that the increased agreement between two data sets does not necessarily make either one correct, since they could agree to 100% and yet still be both wrong. This points to the need for a sustained and continued effort of independent validation of these data which will allow absolute accuracies to be derived for each product and facilitate intercomparisons. Of particular interest to the modeling community will be the validation at coarser resolutions such as $\frac{1}{4}$, $\frac{1}{2}$ or $1°$.

[31] There are a number of issues that remain to be resolved from a continuing dialogue between the users and producers of these global land cover data sets. The first is the need for consistent land cover legends to support a majority of users. This can be facilitated by the development of a global Land Cover Classification System (LCCS) [*DiGregorio and Jansen*, 2000] by the FAO which is a standardized, hierarchical and flexible classification scheme that can be applied irrespective of the source or spatial resolution of the input data. The appeal of the LCCS for global land cover data sets is that is can also be easily collapsed and/or expanded into more or less classes (i.e., cross-walked) to support a wide variety of users. Alternatively, continuous fields approaches that completely bypass the classification scheme may provide a more flexible and accurate approach for land cover product generation. However, this approach will demand some parallel model development so that these data can be used more effectively than they are in current global models. Likewise, classes such as urban areas and wetlands need to be better integrated into future global models.

[32] Finally, the one critical land cover data set that is missing from ISLSCP Initiative II is land cover/use change or disturbance. With new and improved algorithms and data sets, it is now possible to generate time series of land cover products from remotely sensed data, instead of the static, 1992–1993 data set provided here. The challenge remains the integration of "historical" data sets such as those of the AVHRR with those of MODIS and VIIRS, and more importantly, the development of an approach the can consistently separate interannual vegetation changes from actual land cover/use change. It also remains a significant challenge to archive, monitor and generally upkeep the global, high-resolution training data used for the classifications as these change and/or are updated over time. As we look to the future and the NPOESS systems it will be critical that the long-term land cover/use record be established and maintained.

References

Asrar, G., M. Fuchs, E. T. Kanemasu, and J. L. Hatfield (1984), Estimating absorbed photosynthetically active radiation and leaf area index from spectral reflectance in wheat, *Agron. J.*, *76*, 300–306.

Bartholomé, E., and A. S. Belward (2005), GLC2000: A new approach to global land cover mapping from Earth observation data, *Int. J. Remote Sens.*, *26*, 1959–1977.

Bonan, G. B. (1995), Land-atmosphere interactions for climate system models: Coupling biophysical, biogeochemical, and ecosystem dynamical processes, *Remote Sens. Environ.*, *51*, 57–73.

Bonan, G. (1997), Effects of land use on the climate of the United States, *Clim. Change*, *37*, 449–486.

Bonan, G. B. (2002), *Ecological Climatology: Concepts and Applications*, Cambridge Univ. Press, New York.

Bondeau, A., D. W. Kicklighter, J. Kaduk, and the Participants of the Potsdam NPP Model Intercomparison (1999), Comparing global models of terrestrial net primary productivity (NPP): Importance of vegetation structure on seasonal NPP estimates, *Global Change Biol.*, *5*, 35–45.

Bounoua, L., R. S. DeFries, G. J. Collatz, P. Sellers, and H. Khan (2002), Effects of land cover conversion on surface climate, *Clim. Change*, *52*, 29–64.

Brown de Colstoun, E. C., W. Yang, R. DeFries, M. Hansen, and J. Townshend (2000), Surface type visible/infrared imager radiometer suite algorithm theoretical basis document, version 3.0. Natl. Polar-orbiting Oper. Environ. Satell. Syst., Silver Spring, Md. (Available at http://npoesslib.ipo.noaa.gov/u_listcategory_v3.php?35).

Charney, J., P. H. Stone, and W. J. Quirk (1975), Drought in the Sahara: A biogeophysical feedback mechanism, *Science*, *187*, 434–435.

Daughtry, C. S. T., K. P. Gallo, S. N. Goward, S. D. Prince, and W. P. Kustas (1992), Spectral estimates of absorbed radiation and phytomass production in corn and soybean canopies, *Remote Sens. Environ.*, *39*, 141–152.

DeFries, R. S., and S. O. Los (1999), Implications for land-cover misclassification for parameter estimates in global land-surface models: An example from the Simple Biosphere Model (SiB2), *Photogramm. Eng. Remote Sens.*, *65*, 1083–1088.

DeFries, R. S., and J. R. G. Townshend (1994a), NDVI-derived land cover classifications at a global scale, *Int. J. Remote Sens.*, *15*, 3567–3586.

DeFries, R. S., and J. R. G. Townshend (1994b), Global land cover: Comparison of ground-based data sets to classifications with AVHRR data, in *Environmental Remote Sensing From Regional to Global Scales*, edited by G. Foody and P. Curran, pp. 84–110, John Wiley, Hoboken, N. J.

DeFries, R. S., M. Hansen, and J. Townshend (1995), Global discrimination of land cover types from metrics derived from AVHRR pathfinder data, *Remote Sens. Environ.*, *54*, 209–222.

DeFries, R. S., M. Hansen, J. R. G. Townshend, and R. Sohlberg (1998), Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers, *Int. J. Remote Sens.*, *19*, 3141–3168.

DeFries, R. S., J. R. G. Townshend, and M. Hansen (1999), Continuous fields of vegetation characteristics at the global scale at 1 km resolution, *J. Geophys. Res.*, *104*, 16,911–16,925.

DeFries, R. S., M. C. Hansen, J. R. G. Townshend, A. C. Janetos, and T. R. Loveland (2000), A new global 1-km dataset of percentage tree cover derived from remote sensing, *Global Change Biol.*, 6, 247–254.

Dickinson, R. E. (1983), Land surface processes and climate-surface albedos and energy balance, *Adv. Geophys.*, 25, 305–350.

Dickinson, R. E., A. Henderson-Sellers, P. J. Kennedy, and M. F. Wilson (1986), Biosphere-atmosphere transfer scheme (BATS) for the NCAR Community Climate Model, *NCAR Tech. Note NCAR/TN-275+STR*, Natl. Cent. for Atmos. Res., Boulder, Colo.

DiGregorio, A., and L. J. M. Jansen (2000), Land Cover Classification System (LCCS): Classification concepts and user manual, U. N. Food and Agric. Organ., Rome.

Dorman, J. L., and P. J. Sellers (1989), A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model, *J. Appl. Meteorol.*, 28, 833–855.

Eidenshink, J. C., and J. L. Faundeen (1994), The 1 km AVHRR global land data set: First stages in implementation, *Int. J. Remote Sens.*, 15, 3443–3462.

Friedl, M. A., et al. (2002), Global land cover mapping from MODIS: Algorithms and early results, *Remote Sens. Environ.*, 83, 287–302.

Hall, F. G., E. Brown de Colstoun, G. J. Collatz, D. Landis, P. Dirmeyer, A. Betts, G. Huffman, L. Bounoua, and B. Meeson (2006), ISLSCP Initiative II global data sets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies, *J. Geophys. Res.*, doi:10.1029/2006JD007366, in press.

Hansen, M., and B. Reed (2000), A comparison of the IGBP-DIScover and the University of Maryland 1 km global land cover products, *Int. J. Remote Sens.*, 21, 1365–1374.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg (2000), Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, 21, 1331–1364.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, R. Sohlberg, C. Dimiceli, and M. Carroll (2002), Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data, *Remote Sens. Environ.*, 83, 303–319.

Houghton, R. A. (2003), Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000, *Tellus, Ser. B*, 55, 378–390.

James, M. E., and S. N. V. Kalluri (1994), The pathfinder AVHRR land data set: An improved coarse resolution data set for terrestrial monitoring, *Int. J. Remote Sens.*, 15, 3347–3364.

Justice, C. O., J. R. G. Townshend, B. N. Holben, and C. J. Tucker (1985), Analysis of the phenology of global vegetation using meteorological satellite data, *Int. J. Remote Sens.*, 6, 1271–1318.

Kicklighter, D. W., A. Bondeau, A. L. Schloss, J. Kaduk, A. D. McGuire and the Participants of the Potsdam NPP Model Intercomparison (1999), Comparing global models of net primary productivity (NPP): Global pattern and differentiation by major biomes, *Global Change Biol.*, 5, 16–24.

Klein Goldewijk, K. (2001), Estimating global land use change over the past 300 years: The HYDE database, *Global Biogeochem. Cycles*, 15(2), 417–433.

Klein Goldewijk, K., and N. Ramankutty (2004), Land cover change over the last three centuries due to human activities: The availability of new global data sets, *Geojournal*, 61, 335–344.

Loveland, T. R., and A. S. Belward (1997), The IGBP-DIS global 1 km land cover data set, DISCover: First results, *Int. J. Remote Sens.*, 18, 3289–3295.

Loveland, T. R., J. W. Merchant, D. O. Ohlen, and J. F. Brown (1991), Development of a land-cover characteristics database for the conterminous U.S., *Photogramm. Eng. Remote Sens.*, 57, 1453–1463.

Manabe, S. (1969), Climate and the ocean circulation: 1. The atmospheric circulation and the hydrology of the Earth's surface, *Mon. Weather Rev.*, 97, 739–805.

Matthews, E. (1983), Global vegetation and land use: New high-resolution data bases for climate studies, *J. Clim. Appl. Meteorol.*, 22, 474–487.

Melillo, J. M. (1994), Modeling land-atmosphere interactions: A short review, in *Changes in Land Use and Land Cover: A Global Perspective*, edited by W. B. Meyer and B. L. Turner III, pp. 387–410, Cambridge Univ. Press, New York.

Melillo, J. M., et al. (1993), Global climate change and terrestrial net primary production, *Nature*, 363, 234–240.

Mueller-Dombois, D. (1984), Classification and mapping of plant communities: A review with emphasis on tropical vegetation, in *The Role of Terrestrial Vegetation in the Global Carbon Cycle: Measurement by Remote Sensing*, edited by G. M. Woodwell, pp. 19–88, John Wiley, Hoboken, N. J.

Nemani, R., and S. W. Running (1996), Implementation of a hierarchical global vegetation classification in ecosystem function models, *J. Veg. Sci.*, 7, 337–346.

New, M., M. Hulme, and P. Jones (2000), Representing twentieth-century space-time climate variability. Part II: Development of 1901–1996 monthly grids of terrestrial surface climate, *J. Clim.*, 13(13), 2217–2238.

Olson, J. S., J. Watts, and L. Allison (1983), Carbon in live vegetation of major world ecosystems, *Rep. W-7405-ENG-26*, U.S. Dep. of Energy, Oak Ridge Natl. Lab., Oak Ridge, Tenn.

Plattner, G., F. Joos, and T. F. Stocker (2002), Revision of the global carbon budget due to changing air-sea oxygen fluxes, *Global Biogeochem. Cycles*, 16(4), 1096, doi:10.1029/2001GB001746.

Prentice, K. C. (1990), Bioclimatic distribution of vegetation for general circulation models, *J. Geophys. Res.*, 95, 11,811–11,830.

Ramankutty, N., and J. A. Foley (1998), Characterizing patterns of global land use: An analysis of global croplands data, *Global Biogeochem. Cycles*, 12, 667–685.

Ramankutty, N., and J. A. Foley (1999), Estimating historical changes in global land cover: Croplands from 1700 to 1992, *Global Biogeochem. Cycles*, 13, 997–1027.

Running, S. W., and E. R. Hunt Jr. (1993), Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global-scale models, in *Scaling Processes Between Leaf and Landscape Levels*, edited by J. R. Ehleringer and C. Field, pp. 141–158, Elsevier, New York.

Running, S. W., T. R. Loveland, L. L. Pierce, R. R. Nemani, and E. R. Hunt Jr. (1995), A remote sensing based vegetation classification logic for global land cover analysis, *Remote Sens. Environ.*, 51, 39–48.

Scepan, J. (1999), Thematic validation of high-resolution global land-cover data sets, *Photogramm. Eng. Remote Sens.*, 65, 1051–1060.

Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz, and D. A. Randall (1996a), A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: The generation of global fields of terrestrial biophysical parameters from satellite data, *J. Clim.*, 9, 706–737.

Sellers, P. J., B. W. Meeson, J. Closs, J. Collatz, F. Corprew, D. Dazlich, and F. G. Hall (1996b), The ISLSCP Initiative I global datasets: Surface boundary conditions and atmospheric forcings for land-atmosphere studies, *Bull. Am. Meteorol. Soc.*, 77, 1987–2005.

Skole, D. L., and C. J. Tucker (1993), Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988, *Science*, 260, 1905–1910.

Still, C. J., J. A. Berry, G. J. Collatz, and R. S. DeFries (2003), Global distribution of C3 and C4 vegetation: Carbon cycle implications, *Global Biogeochem. Cycles*, 17(1), 1006, doi:10.1029/2001GB001807.

Townshend, J. R. G., C. O. Justice, and V. T. Kalb (1987), Characterization and classification of South American land cover types using satellite data, *Int. J. Remote Sens.*, 8, 1189–1207.

Townshend, J. R. G., C. O. Justice, D. Skole, J.-P. Malingreau, J. Cihlar, P. Teillet, F. Sadowski, and S. Ruttenberg (1994), The 1 km resolution global data set: Needs of the International Geosphere Biosphere Programme, *Int. J. Remote Sens.*, 15, 3417–3442.

Tucker, C. J., J. R. G. Townshend, and T. E. Goff (1985), African land-cover classification using satellite data, *Science*, 227, 369–375.

Wilson, M. F., and A. Henderson-Sellers (1985), A global archive of land cover and soils data for use in general circulation models, *J. Clim.*, 5, 119–143.

———————————

E. C. Brown de Colstoun, Science Systems and Applications, Inc., Code 614.4, Biospheric Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. (ericbdc@ltpmail.gsfc.nasa.gov)

R. S. DeFries and J. R. G. Townshend, Department of Geography, University of Maryland, College Park, MD 20742, USA.